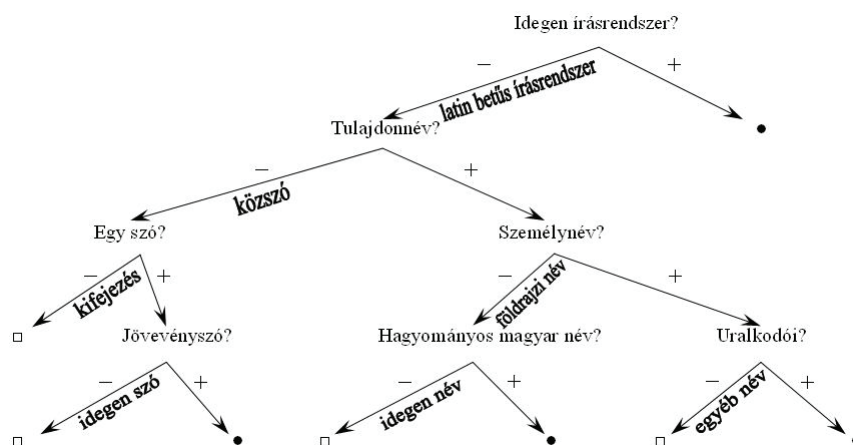


GONDOLATOK A MAGYAR HELYESÍRÁSRÓL EGY ÚJ KÉZIKÖNYV KAPCSÁN II.

Az Osiris Helyesírás kapcsán az idegen szavak és nevek, az átírások, illetve a szótári rész készítésekor felmerült problémákat mutatom be az alábbiakban vázlatosan.

Az **idegen szavak és nevek** írásának központi problémája annak eldöntése, hogy az adott nyelvi alakulatot idegenes vagy magyaros írásmóddal kell-e, célszerű-e rögzíteni. Ennek összefoglalása az alábbi folyamatábrán tehető meg (a □ az idegenes, a ● a magyaros írásmódot jelöli):



Az első fontos különbségtételt az jelenti, hogy egy szó idegen írásrendszerű vagy latin betűs nyelvből származik-e. Az idegen írásrendszerű nyelvekből (pl. héber, japán, kínai, orosz, szerb stb.) átvett szavakat, neveket mindig magyarosan írjuk, többnyire a megfelelő átírási szabályok alapján, néhol hagyományosan kivételes írásmóddal.

A latin betűs nyelvekből átvett közzavak írásmódját az határozza meg, hogy meghonosodott jövevényszóval vagy ún. idegen szóval van-e dolgunk. E két kategória elkülönítésére nem létezik egyértelmű nyelvészeti kritérium, pusztán a nyelvhasználók általános megítélése tekinthető irányadónak. Jövevényszónak az átvett nyelvben meghonosodott, az anyanyelvi beszélők számára már többé-kevésbé felismerhetetlen eredetű szavakat nevezzük, amelyek illeszkedtek az adott nyelv rendszerébe. Ebből következik, hogy a

magyar jövevényszavakat magyarosan kell írni. A helyesírás tehát állást foglal az idegen szó és a jövevényszó megkülönböztetésében: az idegen szavakat idegenes, a jövevényszavakat magyaros írásmóddal kodifikálja. A többszavas közszoí kifejezéseket még akkor is idegenesen kell írni, ha né mely (vagy akár az összes) alkotóelemük külön-külön használva magyarosan íródik (pl. *dzsessz*, de: *hot jazz*).

A latin betűs nyelvekből átvett tulajdonnevek elsősorban földrajzi nevek és személynevek. Ennek a két típusnak a helyesírása részben különbözik. A földrajzi nevek körében jellemzőnek mondható, hogy a magyar kultúrtörténetben fontosabb szerepet játszó helyeknek, különösen településeknek magyar vagy elmagyarosodott nevük van: ezeket értelemszerűen magyarosan írjuk (*Bécs*, *Bukarest*, *Párizs* stb.), bár a nemzetközi használatra szánt kiadványok (térképek, vasúti menetrendek) célszerűen az idegen, helyi névalakkal élnek (*Wien*, *București*, *Paris* stb.). Ha hagyományos magyar névalak nem létezik, az idegen írásmód átvétele a jellemző. Bár régebben a személyneveket szokás volt magyarra fordítani, a lefordítandó nevek köre mára lényegében az uralkodók, illetve a pápák nevére korlátozódik.

A legnagyobb nehézséget az idegen szavak és a jövevényszavak elhatárolása jelenti. Ez a két kategória a nyelvészet eszköztárával ugyanis nem választható szét, a beszélők tudatában mégis külön élnek. A naív nyelvhasználó az előbbiben világosan érzi az idegen eredetet, míg az utóbbiban nem. Mivel azonban minden egyes beszélő kissé másképp ítéli meg a helyzetet, a helyesírás a nyelvközösség általános ítéletét próbálja követni. A két csoportra a következő jegyek jellemzők, de tudni kell, hogy egyik kritérium sem csálthatatlan:

Idegen szavak

- viszonylag frissen kerültek a magyarba, így még nem telt el kellő idő a megszokásukhoz;
- hangtani tulajdonságaik eltérnek a magyarban megszokottaktól;
- valamilyen jellegzetes idegen szóelem található a szóban, rendszerint a szó végén;
- a szó használata valamely csoport- vagy rétegnyelvre korlátozódik.

Jövevényszavak

- régebben kerültek a magyarba, így volt idő megszokni őket;
- hangtani tulajdonságaik összhangban vannak a magyarban megszokottakkal (vagy azért, mert már eredendően is ilyen volt a szó, vagy azért, mert bizonyos hangalaki változás után beilleszkedett a magyar nyelv hangtani rendszerébe);
- nem található a szóban jellegzetes idegen szóelem;
- a szó elterjedt a sztenderdben.

Az **idegen írásrendszerről történő átírással** kapcsolatos első nehézség annak eldöntése, hogy az adott nyelvnek létezik-e latin betűs írásrendszere. Előfordul ugyanis, hogy egy nyelvnek több írásrendszere is van. Ilyen volt a szerbhorvát. Mivel azonban manapság önálló nyelvnek tekintik az egyértelműen latin betűs írású horvát nyelv mellett a szerbet, amelynek írásában egyre kizárólagosabb szerep jut a cirill betűs írásnak, felmerülhet, hogy a korábbi, horvátos átbetűzést idővel a kiejtés szerinti átírás váltja fel. A sajtó gyakorlata ebben a – lehet, elszármazott – áttérésben élen jár, mint általában a helyesírási innovációkban.

Ellentétes irányú folyamat figyelhető meg a japán és a kínai esetében. E nyelvek számára ugyanis nemzetállamaik létrehozta latin betűs írást, és a nemzetközi kommunikációban egyre inkább elvárják ezeknek az átírásoknak a használatát. (Ez a jelenség a globalizációval várhatólag egyre több nyelvet fog érinteni.)

Mivel a nemzetközi átírási rendszerek (egy-egy nyelv esetében gyakran több, általánosan elterjedt, de speciális célú nemzetközi átírási rendszer is él egymás mellett), az átbetűzési rendszerek, illetve az akadémiai helyesírás által kodifikált általános célú magyaros átírások között gyakran van szükség konverzióra, ezt különféle összevető táblázatok segítségével tehetjük meg. Ezek összeállítását és közzétételét igen fontosnak tartottuk az OH. (246–74) elkészítésekor, hogy ezzel is megkönnyítsük a fordítók, a sajtómunkások stb. tevékenységét.

A kötet **szótári részének** jellemzése a következőképpen adható:

Mint minden szótári mű esetében, az egyik legfontosabb jellemzőt a **források** jelentik. Ezek az OH. esetében három fő csoportra tagolódnak: korábbi szótárakra, korábbi szójegyzékekre, digitálisan hozzáférhető modern szövegekre. A korábbi szótárakat egyrészt értelmező, másrészt helyesírási szótárak reprezentálják, a szójegyzékek is két fő csoportba, a helyesírási és a tematikus szójegyzékek csoportjába sorolhatók.

A szótári rész számára az új szavak, szókapcsolatok gyűjtése digitálisan hozzáférhető modern szövegek feldolgozásával történt a különféle, tapintható hiányokat pótlandó. Mivel a tágran értelmezett köznyelv és hivatalos nyelv szókészlete eleve alaposan fel volt dolgozva a Magyar értelmező kéziszótár, illetve a Magyar helyesírási szótár (és előzményei) anyagában, az általános műveltség szakszó- és tulajdonnévanyagára különös tekintettel volt a Helyesírási diákszótár, a leginkább érzékelhető szókészletbeli hiányok pótlására az alábbi szövegtípusok látszottak leginkább megfelelőnek: 1. modern tudományos-ismeretterjesztő szövegek; 2. mai sajtónyelv, elsősorban publicisztika; 3. hagyományos női terepek szókészlete.

Az így gyűjtött több tízmillió szövegszóra rúgó forrásszöveg-gyűjtéményt számítógépes módszerekkel dolgoztuk fel. Első feladat volt a hasznos és a haszontalan szövegrészek – nagyban automatizált – szétválasztása (jellegzetes példa erre az egyes elektronikus sajtótermékek rendszeresen ismétlődő fejrészei, amelyek a későbbi statisztikai feldolgozást nagymértékben torzították volna). A következő lépést a szóalakok gyakoriság alapján történő gépi gyűjtése jelentette, de tekintettel kellett lenni a szavakon kívül a tipikus együttes előfordulásokra is. Ebben felhasználtuk a modern kollokációkutatás eredményeit.

Az informatikai szempontból talán legizgalmasabb feladatot az jelentette, hogy ebből a szóalakjegyzékből a szótári szók jegyzéke minél kisebb emberi munkával előállítható legyen. A számítógépes morfológiai elemzéshez általában szükség van tőtárra, mivel azonban jelen esetben épp a tőtár bővítése a feladat, ez az út nem járható. Ily módon a lehetséges megoldás az ún. heurisztikus tövesítésben rejlett, azaz ha a szóalakjegyzéket feldolgozó program bemenetén megjelent például a *plug-in* és a *plug-inek* szóalak egyaránt, akkor az utóbbi eldobható, ugyanis szabállyal jól leírható, morfoszintaktikailag transzparens toldalékról van szó. Ezzel a módszerrel a manuálisan átvizsgálandó szóalakok száma radikálisan csökkenthető volt. Az így kiszűrt kb. kétszázezer szóalak és szókapcsolat emberi erővel történő átfésülése mintegy hússzezer új szótárérett szó és szókapcsolat felvételét tette lehetővé.

Egy szótár másik fontos jellemzőjét a **menyiségi tulajdonságok** írják le. Az OH. szótári része 47 186 szócikkben 213 974 (illetve önállóan számítva a kiejtést, elválasztást, ragozott alakokat, több mint háromezrezer) szótári tételt dolgoz fel, ezzel jelentősen meghaladva minden korábbi köznyelvi szótárat.

Az **idegen szavak** kezelése nemcsak a tanácsadó részben, hanem a szótárban is felveti az idegen vs. magyaros írásmód kérdését. Mivel ez számos egyedi döntést igényelne, célszerűnek látszott a következő irányelvek betartása:

- Amennyiben korábbi helyesírási szótárakban már megjelent a magyaros írásmód, az írásgyakorlattól függetlenül feltétlenül a magyaros írásmódot követtük, így az úzusban egyértelműen erősebb jazz, grapefruit helyett dzsessz, grépfrút szerepel a szótárban. (Fontos kitétel ebben az irányelvben a korábbi szótárak helyesírási volta, hiszen az értelmező, illetve idegenszó-szótári hagyomány esetenként komoly eltéréseket mutathat ettől a műfajtól.)
- Amennyiben korábbi helyesírási szótárakban csak idegenes írásmód szerepel, rendszerint az idegenes írásmódot követtük az OH.-ban is. Kivételt jelentenek ez alól a Magyar helyesírási szótár 1999-es megjelenése óta magyarosan erősen elterjedt szavak, mint például a dizájn.

- Amennyiben korábbi helyesírási szótárban elő nem forduló szóról, kifejezésről kellett állást foglalni, a Magyar nemzeti szövegtár előfordulási gyakoriságát vettük figyelembe. Ha legalább 40% volt a magyaros írásmód aránya, rendszerint a magyaros írásmódot láttuk célszerűnek követni. Itt természetesen nem lehetett szó mechanikus döntésről, hiszen bizonyos rendszerszerűségek felülírhatták az így adódó döntést, illetve figyelemmel kellett arra is lenni, hogy ha lényegében csak a személyes alkorpuszban jelent meg innovációként a magyaros írásmód, az nem lehet elégséges ennek kodifikálásához.

Az ezredforduló komoly lexikográfiai projektje a **szótárak határtalanítása**, azaz annak megvalósítása, hogy a magyar szótárak ne csupán a magyarországi magyar nyelvet mutassák be (vö. Lanstyák 2005). Ennek nyilvánvalóan nagyobb tere van egy deskriptív szótárban, mint egy műfajából adódóan bizonyos mértékig kényszerűen preskriptív helyesírási kézikönyvben. Ezért érdemes néhány összevető megjegyzést tenni, hogy miben más egy határtalanított helyesírási szótár az első határtalanított magyar szótárhoz, a Pusztai Ferenc főszerkesztésében megújított Magyar értelmező kéziszótárhoz (2003) képest.

Az OH. újdonsága, többlete, hogy minden Kárpát-medencei terület (Erdély, Felvidék, Kárpátalja, Vajdaság, Drávaszög, Órvidék, Muravidék) szókészletéből válogat, hiszen mostanra már elkészültek a vajdasági, drávaszögi, órvidéki és muravidéki szólisták is (az utóbbi szókészletéről részletesen szól Kolláth 2005). Helyesírási szótárról lévén szó, azonban az ÉKsz.²-höz képest bizonyos szempontok mentén szűkíteni is kellett a bemutatott szókészletet: kitartottunk a mellett az általános elvünk mellett, hogy csak írott nyelvből is jól adatolható szavakat, kifejezéseket szótárazunk, nem akarván állást foglalni abban a kérdésben, hogy írásbeliséggel gyakorlatilag nem rendelkező szavak esetében az idegenes vagy a magyaros írásmód preferálandó-e. Ez az elv természetesen maga után vonja, hogy a hivatalos nyelv szókészletének elemei felülreprezentálódnak a mindennapi beszélt nyelvéhez képest (ezért van például álláspontunk szerint helye egy helyesírási szótárban az *alapiskolának*, szemben a *tyepivel*). További – műfajból adódó – bővülés a helyesírási szótárban létjogosult tulajdonnevek (jelen esetben elsősorban intézménynevek) szerepeltetése, ami viszont elképzelhetetlen értelmező szótárban. (NB. Nem csekély feladat a természetföldrajzi névanyag újragondolása sem, hiszen a jelenleg általánosan használt, per definitionem országhatároktól független természetföldrajzi tereptárgyak elnevezései implicite utalnak a mai országhatárokra; vö. pl. *Északi-középhegység*. Erre tesz kísérletet többek között Hajdú-Moharos–Hevesi 1997. Vö. még Faragó 2001.)

Mivel az OH. szótári része számos tömörítő eljárást alkalmaz, ez kötöttebb, de némiképp szokatlan szócikkstruktúrát eredményez, végezetül ezt mutatom be egy rövid mutatványszócikken:

báró bárót, bárók, bárója || báróné, báróság | bárónő | báró Eötvös [ötvös] {Eöt-vös}
József ~ Eötvös József báró || → drog, ipar, zöld | Coubertin, Eötvös,
Fejérváry, Fourier, Korányi, Macaulay, Podmaniczky, Wesselényi

A félkövérrel szedett címszót követik a szófajra legjellemzőbb, a paradigmát bemutató ragos alakok. A szócikk alapegységeit kettős virgula választja el, így a szócikkfejet követő fő részt is ez vezeti be. A fő részben először a képzős származékok, majd puszta virgula után a címszóval mint előtaggal alkotott összetételek, végül újabb puszta virgulát követően a címszóval mint vezérszóval, meghatározó taggal alkotott szókapcsolatok. A szócikknek eddig a részéig fordulhatnak elő kiejtési, illetve elválasztási információk szögletes, illetve kapcsos zárójellel jelölve. A szócikk végén megjelenő utalókat ismét kettős virgula vezeti be, ezt követi az utalást szimbolizáló nyíl. Először azon címszavak következnek, amelyek szócikkében összetételi utótagként fordul elő a szó, majd puszta virgula után azok a címszavak, amelyek szócikkében szókapcsolat alaptagjaként vagy egyéb nem vezérszói pozícióban.

Ez az elv olyan – első látásra mehökkentő – utalósort is eredményezhet, mint amilyen például az **aktív** szócikkben szerepel: *bioaktív, felület, hiper-* stb. Az utalózott szóadatok ebben az esetben: *bioaktív, felületaktív, hiperaktív*. Ez a látszólag következtelen eljárás amiatt van, hogy a *bioaktív* (szókapcsolata miatt) önálló szócikkben fordul elő, a *felületaktív* a **felület** címszó összetételi bokrában, míg a *hiperaktív* a **hiper-** szócikkben: a kötőjel arra utal, hogy önálló szóként a magyar sztenderdben **hiper* szó nem létezik, csak előtagként.

A szókapcsolati utalók példája jól illusztrálja, hogy helyesírási információkon kívül bizonyos esetekben műveltségbeli tényanyagot is be kíván mutatni a szótár, így egy-egy nevezetesebb névnap alkalmából például az iskolában könnyen kikereshető a híres Jánosok vagy Máriák. Ugyanígy a szókapcsolati utalók segítségével egy helyen olvasható az összes történelmi vármegye vagy mai kistérség neve.

Szakirodalom

Faragó Imre 2001. A magyar névhasználat változásai a Kárpát-medencét ábrázoló térképeken II. Geodézia és Kartográfia 53/2.

- Hajdú-Moharos József – Hevesi Attila 1997. A kárpát-pannon térség tájtagolódása.
In: Karátson Dávid (szerk.): Pannon enciklopédia. Magyarország földje.
Kaloprint, Kalocsa.
- Kolláth Anna 2005. Magyarul a Muravidéken. Zora 39. Slavistično Društvo, Maribor.
- Laczkó Krisztina – Mártonfi Attila 2004. Helyesírás. Osiris Kiadó, Budapest.
- Lanstyák István 2005. Határtalanítás. A Magyar értelmező kéziszótár 2. kiadása
után, 3. kiadása előtt. In: Mártonfi Attila – Papp Kornélia – Slíz Mariann
(szerk.): 101 írás Pusztai Ferenc tiszteletére. Argumentum Kiadó, Budapest.
181–8.
- Pusztai Ferenc (szerk.) 2003². Magyar értelmező kéziszótár. Akadémiai Kiadó,
Budapest.